AAMRL-TR-85-046

BANDPASS SPATIAL FILTERING AND INFORMATION CONTENT

DTIC
S ELECTE D
JAN 03 1996
F

GILBERT G. KUPERMAN

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY

JULY 1985

19951228 048

Approved for public release; distribution unlimited.

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

DTIC QUALITY INSPECTED 1

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Armstrong Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-85-046

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified

1b. RESTRICTIVE MARKINGS:

2a. SECURITY CLASSIFICATION AUTHORITY:

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE:

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S): AAMRL-TR-85-046

5. MONITORING ORGANIZATION REPORT NUMBER(S):

6a. NAME OF PERFORMING ORGANIZATION: Harry G. Armstrong Aerospace Medical Research Laboratory

6b. OFFICE SYMBOL *(If applicable)*: HED

7a. NAME OF MONITORING ORGANIZATION:

6c. ADDRESS *(City, State and ZIP Code)*: Wright-Patterson AFB, Ohio 45433-6573

7b. ADDRESS *(City, State and ZIP Code)*:

8a. NAME OF FUNDING/SPONSORING ORGANIZATION:

8b. OFFICE SYMBOL *(If applicable)*:

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER:

8c. ADDRESS *(City, State and ZIP Code)*:

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
62202F	7184	11	44

11. TITLE *(Include Security Classification)*: BANDPASS SPATIAL FILTERING AND INFORMATION CONTENT (U)

12. PERSONAL AUTHOR(S): Kuperman, Gilbert G.

13a. TYPE OF REPORT: Final

13b. TIME COVERED: FROM 3-1-85 TO 6-4-85

14. DATE OF REPORT *(Yr., Mo., Day)*: 1985 July

15. PAGE COUNT: 50

16. SUPPLEMENTARY NOTATION:

17. COSATI CODES

FIELD	GROUP	SUB. GR.
05	05	
05	08	

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*: Spatial Filtering; Perception; Image Quality Assessment; Digital Image Processing

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Gaussian weighted, bandpass spatial filters were applied to pictures of tactical fighter aircraft through two-dimensional, Fourier Transform processing. Subjects were required to rate image interpretability and report their confidence. No significant difference was found between aircraft for interpretability or confidence. Filter center frequency was found to be highly significant ($p < 0.01$), and the difference between subjects was found to be significant ($p < 0.05$) with respect to interpretability. The main effect of filters was found to be significant ($p < 0.05$) for the dependent measure of confidence. A predictor equation was developed through linear regression to relate filter center frequency and image interpretability ($r^2 = 0.5623$).

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☒ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL: Gilbert G. Kuperman

22b. TELEPHONE NUMBER *(Include Area Code)*: (513) 255-3727

22c. OFFICE SYMBOL: AAMRL/HED

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

PREFACE

This work was performed in the Visual Display Systems Branch of the Harry G. Armstrong Aerospace Medical Research Laboratory's Human Engineering Division (AAMRL/HE). The effort was in support of Work Unit 7184 11 44, "Image Display Mensuration/Enhancement."

The author wishes to thank Mr. R. Kevin Burns (ASD/ENECH) who, as part of a practicum associated with a Masters Degree program in Applied Behavioral Science, assisted in the imagery digitization and preparation of other study materials and Ms. Denise Wilson (SAIC) who assisted in the experimental design and analysis. Special thanks are due to the management and personnel of the Sensor Evaluation Branch of the Air Force Wright Aeronautical Laboratories (AFWAL/AARF) who participated in the study.

Accesion For

NTIS CRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification

By
Distribution /

Availability Codes

Dist	Avail and / or Special
A-1	

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Section 1
INTRODUCTION

The Air Force has demonstrated a long term and continuing interest in the abilities of the human observer to abstract information from imagery displays. This interest has manifested itself in two primary areas:

- The interpretation of reconnaissance/surveillance imagery for intelligence and targeting purposes.

- The acquisition of targets for weapon delivery.

In the first case, imagery is exploited to satisfy stated essential elements of information (EEIs) that serve as part of a broader intelligence collection effort. These EEIs may be specific to the imagery acquisition mission as, for example, "Are there tanks at coordinates XXXX N YYYY E?" In other cases, the EEIs may be in the form of standard information sets to be provided. An Air Reconnaissance Target Reporting Guide (Anonymous, 1969), for example, calls for the following information to be provided from imagery of airfields:

- Serviceability.
- Type and location of aircraft present.
- General activity (special purpose, nonaircraft).
- Runway pattern, dimensions, and surface materials.
- Types, numbers, and locations of defenses.
- Type, number, size, and location of fuel storage.
- Number and location of ammunition storage facilities.
- Number, type, construction, and sizes of hangars.
- Number, type, and location of electronic facilities.

Reconnaissance imagery exploitation is usually not severely time critical; and cycle times, from receipt of imagery at the interpretation station to the production of an interpretation report, typically take from several minutes to hours.

In the target acquisition case, two purposes are to be satisfied. First, the pilot (or radar navigator, offensive systems operator, weapon systems operator, etc.) must confirm that the object imaged is the target of interest. Second, he must designate, from the displayed image, the target's location to a weapon delivery computer. These tasks are much more time-compressed, with timelines of only several seconds being typical. Because of the time constraint, great reliance is placed on real-time and near real-time sensors [e.g., forward-looking infrared (FLIR) and synthetic aperture radar (SAR)].

A systems approach to these cases decomposes the problem into three subsystems: the sensor, the display medium, and the operator. Research and development activities continue in each of these areas. Kuperman et al. (1977) present a series of studies and analyses addressing the selection of near-real-time and real-time sensors for a remotely piloted vehicle tasked with a reconnaissance mission. Task (1979) provides a review and empirical comparison of several methods and models applicable to assessing the image quality of video displays. Research to quantify the abilities of the observer range from basic studies of visual perception (e.g., Carlson and Cohen, 1978) through highly applied design information (e.g., Erickson, 1978).

A related area of research has attempted to deal with the information content and utility of the displayed information to specific task requirements. Johnson (1958) established the concept of equating system resolution (lines across the target's critical or minimum dimension) with performance in reporting the detection, classification, and recognition of military targets. Nygaard et al. (1964) attempted to relate a measure of information content that they termed "stimulus complexity" to observer performance. More recently, researchers have attempted to construct and apply psychovisually based models of perception to the information extraction task. Campbell et al. (1970), among many others, reported on the ability of the human to distinguish between gratings of sinusoidally modulated luminance at different spatial frequencies. The explicit application of the techniques of Fourier decomposition and synthesis were reported in Campbell et al. (1968). Parallel applications of Fourier techniques, specifically the

Modulation Transfer Function (MTF), have taken place in the evaluation of sensor image quality (e.g., Gliatti, 1977; and Kuperman, 1980) and in creating visual describing functions (e.g., Cornsweet, 1970).

One approach to applying Fourier transform based methods to the creation of visual system describing functions being pursued at Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL) was reported by Ginsburg (1978). He applied the Contrast Sensitivity Function (CSF) to attempt to account for a wide range of visual phenomena including illusions, texture perception, and object recognition. The CSF is essentially a visual system demand curve that describes the amount of modulation required for detection as a function of spatial frequency. Ginsburg suggested that evidence existed that the CSF was actually the envelope of a family of Gaussian weighted, bandpass spatial filters separated from each other in center frequency at 1 octave intervals and each exhibited at 1 to 2 octave bandwidth. [Two dimensional, discrete Fourier transform applications in digital image processing are described in Lewis (1984)]. In relating the CSF to the identification of objects, he stated:

> "Each task and each object requires slightly different numbers of spatial frequencies or harmonics. For example, to detect the presence of an object requires only very low spatial frequency, less than the fundamental spatial frequency of the object. The classification of Snellen letters requires from about 1.5 to 2.5 cycles per letter. The identification of a face from a small class of faces requires about 4 cycles per face width."

The CSF exhibits a peak response at approximately 2 cycles per degree and is essentially zero at spatial frequencies greater than 65 cycles per degree. (In order to permit generalization to varied imaging and viewing conditions, the resolution of the acquired imagery is specified in object domain (i.e., cycles per target critical dimension, while the visual requirements of the observer are stated in terms of the number of cycles per angular subtense at the eye.) The family of spatial filters is also referred to as a channel model of perception.

Other research (Snyder, 1974) has attempted to form a unitary measure of image utility, including the observer, by bounding the area below the system

MTF curve with an observer visual threshold curve and reporting the area contained between the two. The channel model representation of the CSF suggests that visual perception can be considered to be a hierarchial decomposition of visual stimuli, with each channel (filter) making distinct contributions to the percept. Johnson (1958) reported that increasing levels of information (cycles per target dimension) were required to support increasing levels of information extraction. These findings suggested that the bandpass filters should demonstrate a relationship with an image quality measure based on distinct level of information extraction. The present research is a pilot study seeking to investigate this relationship.

Section 2
INFORMATION EXTRACTION

Numerous approaches have been attempted to define the information content levels required to support different levels of observer task performance. The Johnson Criteria (1958) state that target detection requires 1.0 ± 0.25 cycles per object dimension, orientation reporting requires 1.4 ± 0.35 cycles, target recognition requires 4.0 ± 0.8 cycles, and target identification requires 6.4 ± 1.5 cycles. This early work has been refined by others (for example, Erickson, 1978) but has remained reasonably unchanged over time.

A standard imagery interpretation handbook, AFM 200-50 (1967), attempts to provide guidelines for specifying the image scale required in aerial photography to support target identification and detail interpretation. The following table is extracted for aircraft.

TABLE 1. MINIMUM SCALES FOR INTERPRETATION AND IDENTIFICATION

Wing Span	Minimum Scale for Interpretation	Minimum Scale for Detail Interpretation
Under 40 feet	1/10,000	1/2,000
40 to 60 feet	1/12,000	1/3,000

These scales are based on the assumption that the photographic sensor exhibited an average system resolution of 15 to 20 line pairs per millimeter. Assuming 20 lines per millimeter resolution, the required scale factors can be converted into required ground resolved distances (GRDs).

TABLE 2. GRDs FOR INTERPRETATION AND IDENTIFICATION

Wing Span	GRD for Interpretation	GRD for Detail Interpretation
Under 40 feet	6.6 feet (2.0 m)	1.3 feet (0.4 m)
40 to 60 feet	7.9 feet (2.1 m)	2.0 feet (0.6 m)

An Image Interpretability Rating Scale was adopted (ASCC, 1978) by the representatives of the U.S. Air Force, U.S. Navy, Royal Air Force, Royal Australian Air Force, Royal New Zealand Air Force, and Canadian Forces for the reporting of photographic image quality. This scale provides 10 rating categories, 0 through 9. Category 0 is described as being useless for interpretation due to cloud cover, poor resolution, etc. Category 4 would support the recognition of aircraft, for example, when singly deployed. The GRDs that correspond to Category 4 are in the range of 3.9 feet (1.2 m) through 8.2 feet (2.5 m). Category 6 of this scale pertains to imagery of sufficient quality to support the identification of singly deployed aircraft. The GRD range for this level is 1.3 feet (0.4 m) through 2.5 feet (0.75 m). Snyder et al. (1982) reported on the application of this rating scale to the image quality assessment of digitally processed, hardcopy imagery that had been degraded by various levels of noise and blur. The ratings were then used to predict the ability of military photointerpreters to perform an information extraction task: the response to a set of EEIs established for each scene. They found that "mean scale scores correlated quite well [r = 0.90] with information extration performance." [Ward et al. (1984) report on a similar type of rating scale developed for synthetic aperture radar sensors.]

Section 3
METHODOLOGY

STIMULI

Stimulus preparation was composed of two steps. First, a set of line drawings of tactical fighter aircraft was digitized. Second, digital filtering was applied to these images to create the stimuli conditions investigated in the study.

An image interpretation aid, a U.S. fighter aircraft comparison wheel, produced by the 460th Reconnaissance Tactical Squadron, Langley Air Force Base, served as the source material. This wheel consists of a transparent overlay that is free to rotate around its center. The underlay is printed hardcopy. Both the overlay and hardcopy contain line drawings of 12 aircraft. Six of these were selected for use in the study. They were the A-6, A-7, F-4, F-14, F-15, and F-16. All are currently flying with the USAF or USN. The dimensions of the aircraft are presented in Table 3.

TABLE 3. AIRCRAFT DIMENSIONS

Aircraft	Length	Width
A-6	54' 7"	53' 0"
A-7	46' 1.5"	38' 9"
F-4	62' 11.74"	38' 5"
F-14	61' 10.6"	64' 1.5" (unswept)
F-15	63' 9.75"	42' 9.75"
F-16	47' 0"	30' 0"

The line drawings of these from the transparency portion of the comparison wheel were digitized in the Visual Image Processing Enhancement and Reconstruction (VIPER) facility of AAMRL. VIPER is a digital image processing facility dedicated to supporting research in human engineering (Kuperman, et al., 1984).

Digitization of the line drawings was performed using a video digitizing capability. Each aircraft image was captured in a 256 by 256 picture element (pixel) array using 8 bits (256 levels) of intensity code. The working distance of the digitizing camera was adjusted so that the average of the aircraft dimensions [(length + width)/2] was approximately half the size of the digital array into which each picture was captured. (The average of the measured dimensions for all six aircraft was found to be 126.7 pixels.) The six images were edited to remove nonhomogeneous areas in their respective backgrounds. Figure 1 presents reconstructions of the six aircraft images. The first column in the figure (from top to bottom) contains the A-6, A-7, and F-4 aircraft. The second column contains the F-14, F-15, and F-16.

Once the aircraft drawings had been digitized, filtered versions of each aircraft had to be created The specific filters employed were derived from the channel model of visual perception. Ginsburg (1980) describes the filters as being 1 octave apart and having bandwidths of 1 to 2 octaves. Thus, if a filter has a center frequency of 1 cycle, then the next filter would have its center frequency at 2 cycles (i.e., 1 octave). If the filter centered at 2 cycles was 1 octave wide in bandwidth, then its lower frequency cutoff would be at 1 cycle and its upper frequency cutoff would be at 4 cycles. For this study, a bandwidth of 1.5 octaves was selected. The center frequencies were 0.5, 1.0, 2.0, 4.0, 8.0, 16.0, and 32.0 cycles per digital image dimension. Table 4 presents the passbands of these filters.

TABLE 4. GAUSSIAN FILTER PASSBANDS

Lowest Frequency	Center Frequency	Highest Frequency
0.0	0.5	0.8
0.0	1.0	1.7
0.8	2.0	3.4
0.4	4.0	6.7
0.2	8.0	13.5
0.1	16.0	26.9
0.0	32.0	53.8

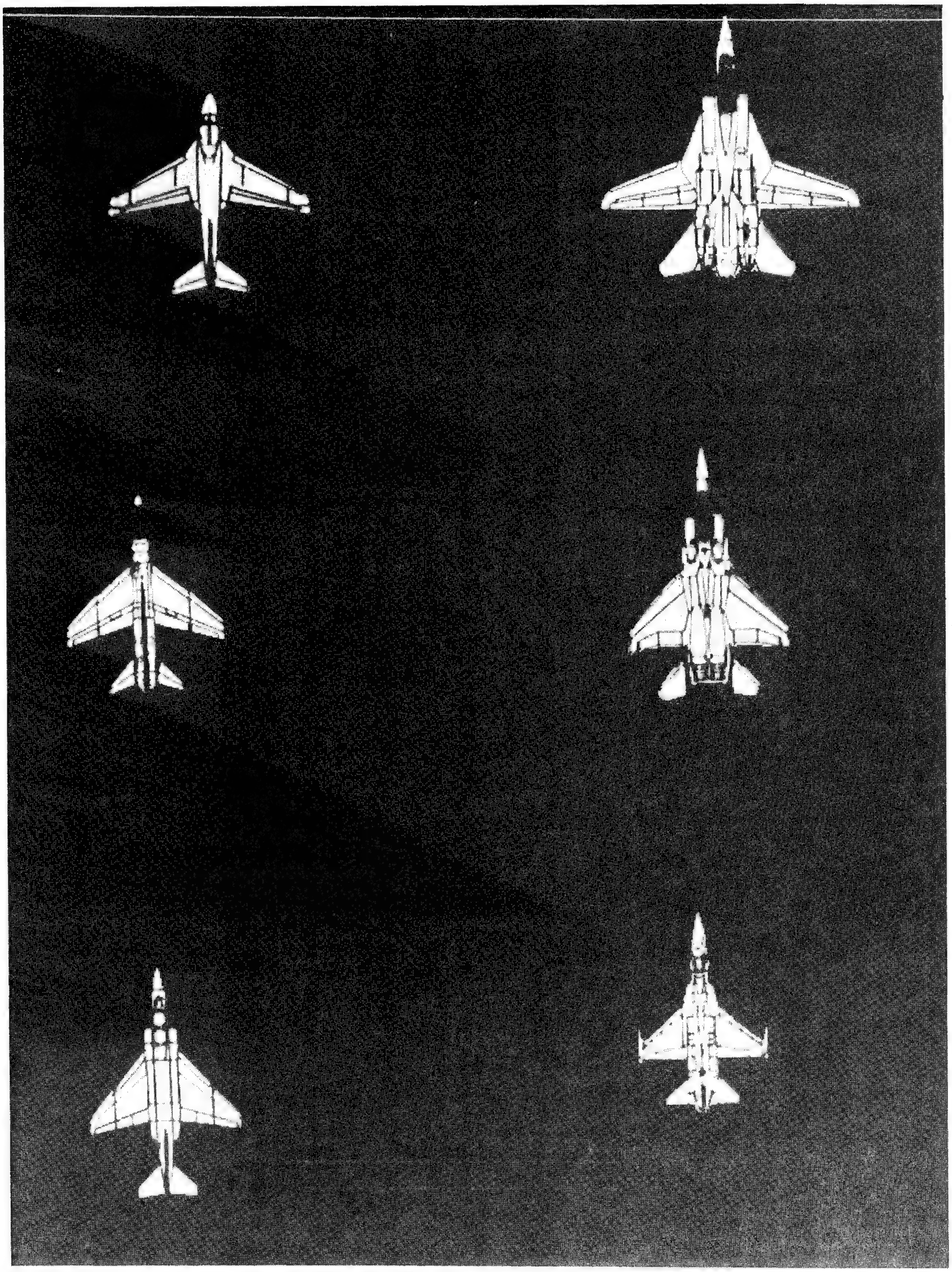

Figure 1. Digitized Versions of Six U.S. Fighter Aircraft

Since the average dimension of the six aircraft was approximately half the size of the total digital image, the center frequencies corresponded to 0.25, 0.5, 1.0, 2.0, 4.0, 8.0, and 16.0 cycles per object dimension. Remembering that one cycle is approximately equivalent to two lines of resolution, the range of center frequencies fully spans the resolution range reported to be required for aircraft recognition. (Erickson, 1978, cited research results which stated that approximately 12 lines per object produced the most accurate responses in an aircraft identification task.)

Figure 2 presents a visualization of the seven bandpass filters. In the figure, the response of the filter (i.e., what proportion of the signal contained in the Fourier transform, at a specific spatial frequency and orientation) is depicted in terms of relative reflectance. Thus, the brighter portions of the filter, at and near to its center frequency, pass spatial frequencies to a greater extent than do the darker areas (those further away from the center frequency). It should also be noted that the filters are two dimensional, being circularly symmetric about the zero spatial frequency (or D.C.) point in the transform plane.

Filtered versions of each the six aircraft images were created for each channel model filter. Figure 3 depicts the digital image processing flow followed in creating these filtered versions.

First, the seven filters were created. A VIPER image processing command, two-dimensional filter (TDF) was employed. TDF requires the user to specify the shape of the filter, its center frequency, and bandwidth. In this case, these options were Gaussian filter, center frequencies of 0.5, 1.0, 2.0, 4.0, 8.0, 16.0, and 32.0 cycles per image dimension, and a bandwidth of 1.5 octaves. Thus, a set of seven filters was created through repeated invocation of the TDF command. Second, the image of each aircraft was subjected to Fourier transformation (FT). Seven copies of each transformed aircraft image were made, one to correspond to each filter. Next, each copy of the FT of the aircraft was multiplied by its respective filter. Last, each filtered aircraft image transform was subjected to the inverse FT process (FT^{-1}). This returned the filtered images from the transform plane to the object plane. A total of 42 filtered images were produced in this

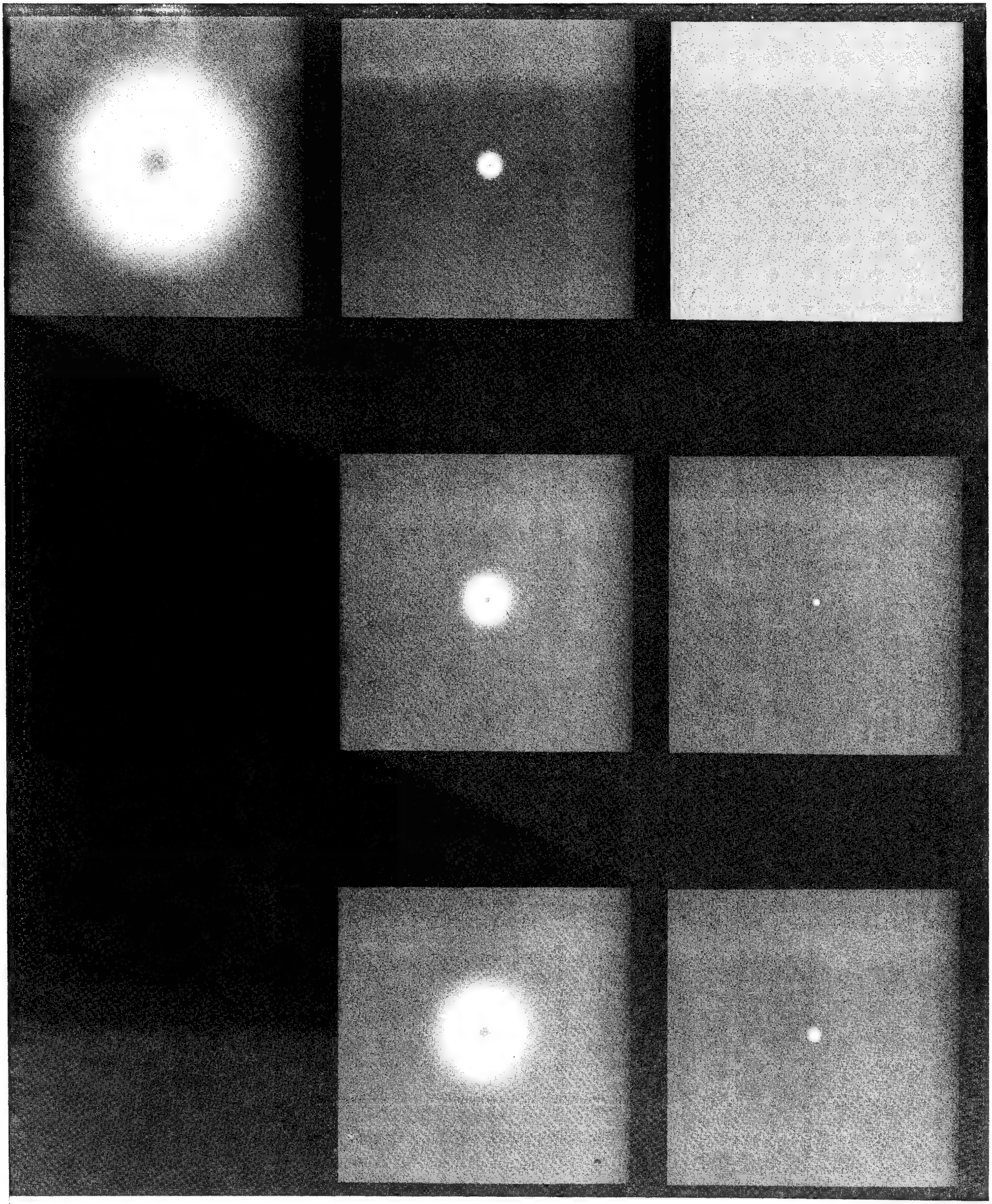

Figure 2. Reconstructions of Channel Model Gaussian Bandpass Spatial Filters

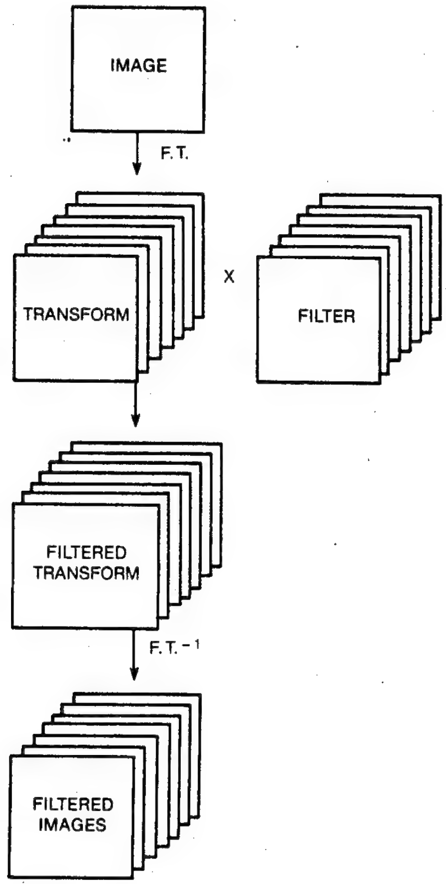

Figure 3. Digital Filtering Flow Diagram

manner (six aircraft x seven filters = 42 filtered images). The six original aircraft images and the 42 filtered versions were recorded on Type 52 Polaroid print film (as positive prints) using a Matrix Instruments Color Graphics Camera. Each picture (aircraft and filtered version) was a square, approximately 1.5 inches (38 mm) on each edge.

Figures 4 through 9 present the filtered image series for each aircraft. In each of these figures, the filtered images are presented, from left to right and top to bottom, in order of the (increasing) center frequency of the filter used in creating them. In each figure, the Fourier transform is presented, in image form, in the lower right hand corner.

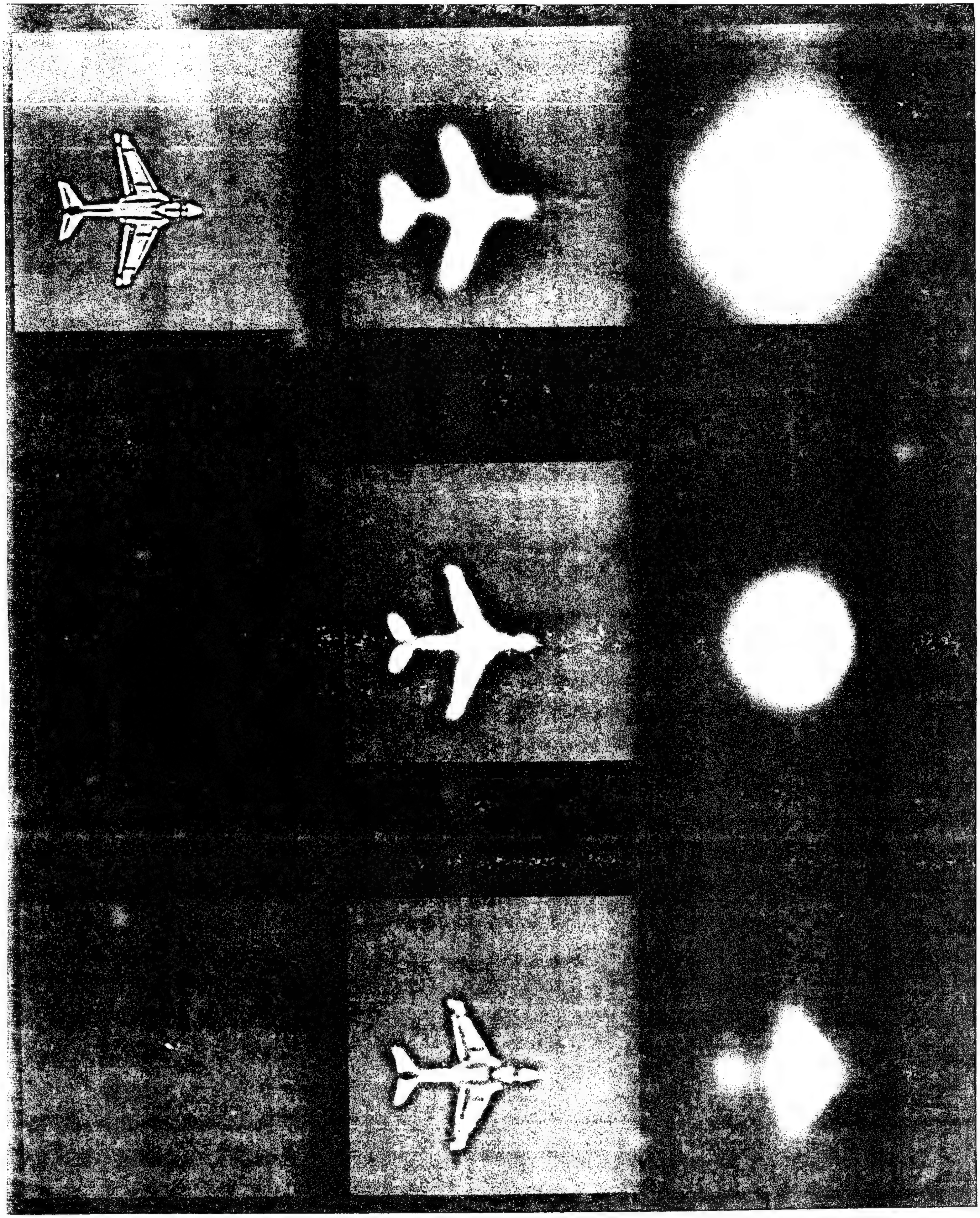

Figure 4. Bandpass Spatial Filtered Series, A-6 Aircraft

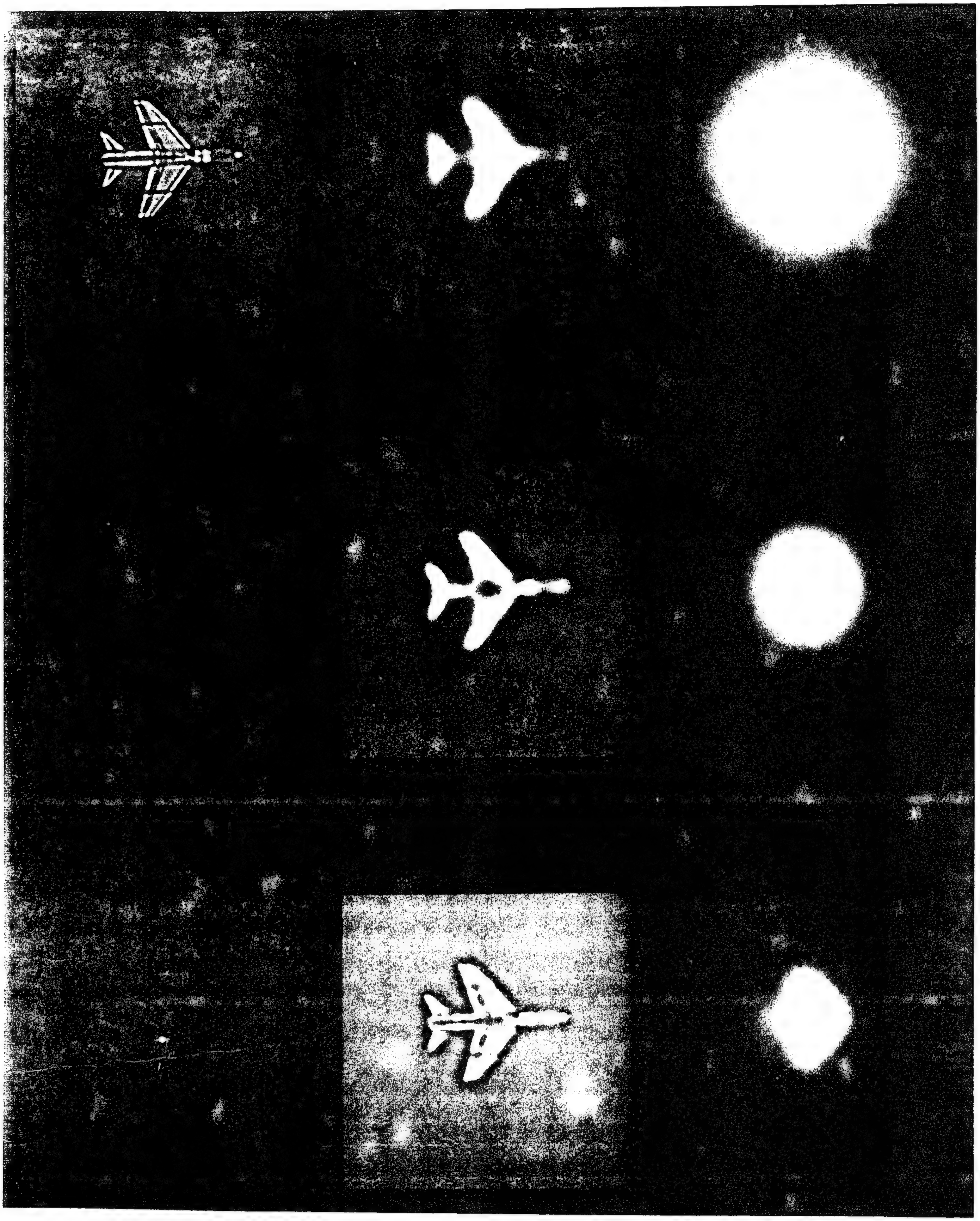

Figure 5. Bandpass Spatial Filtered Series, A-7 Fighter

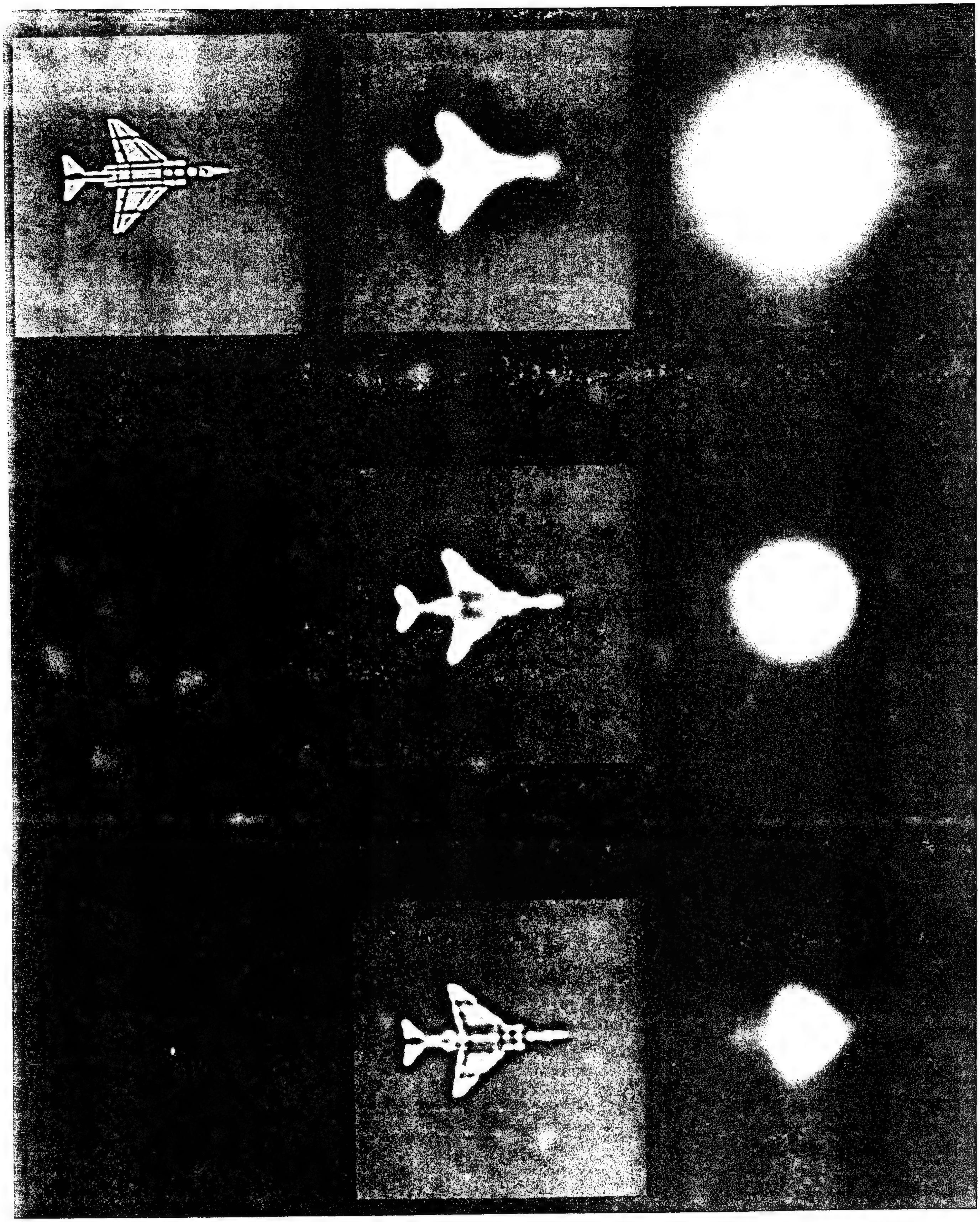

Figure 6. Bandpass Spatial Filtered Series, F-4 Fighter

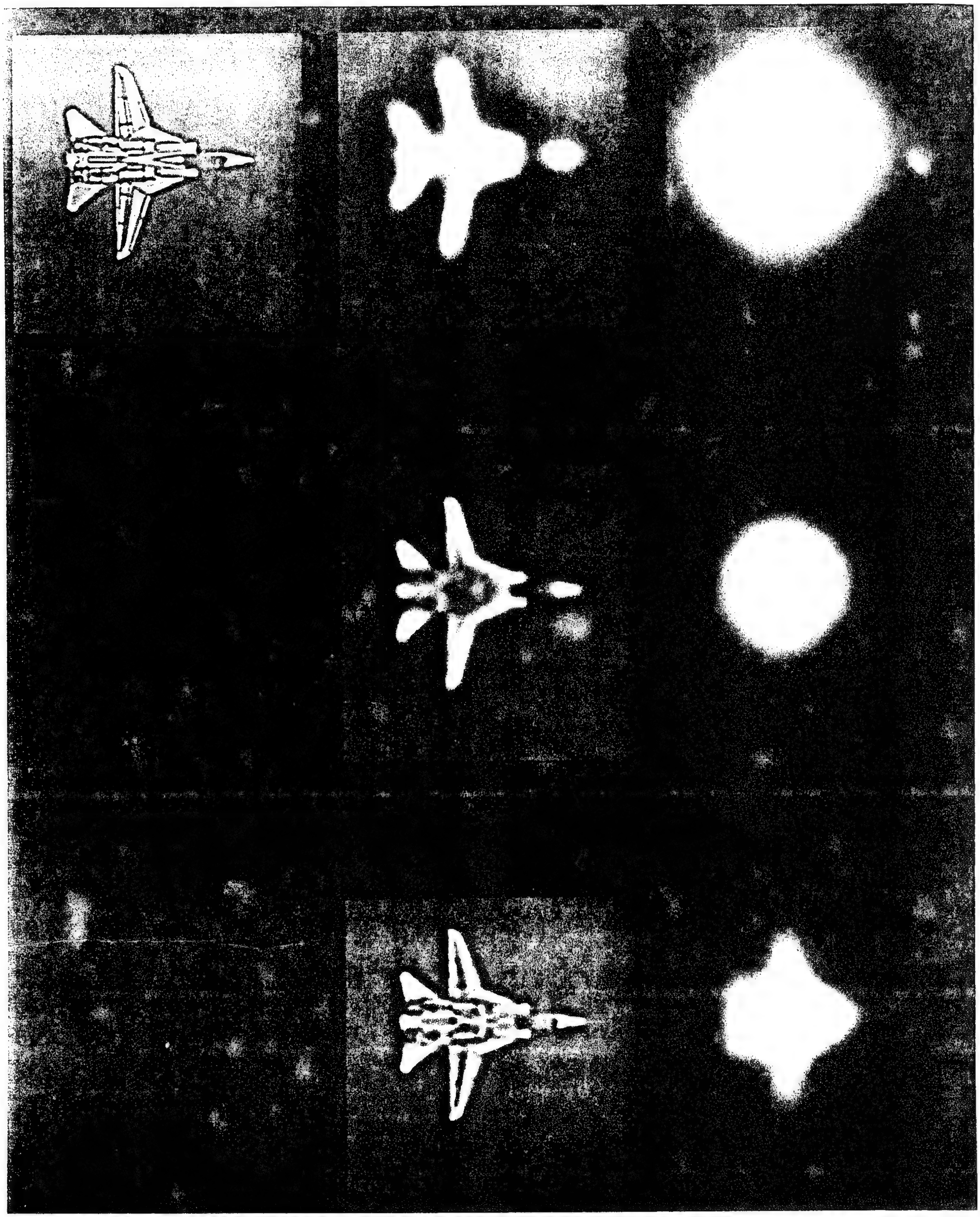

Figure 7. Bandpass Spatial Filtered Series, F-14 Fighter

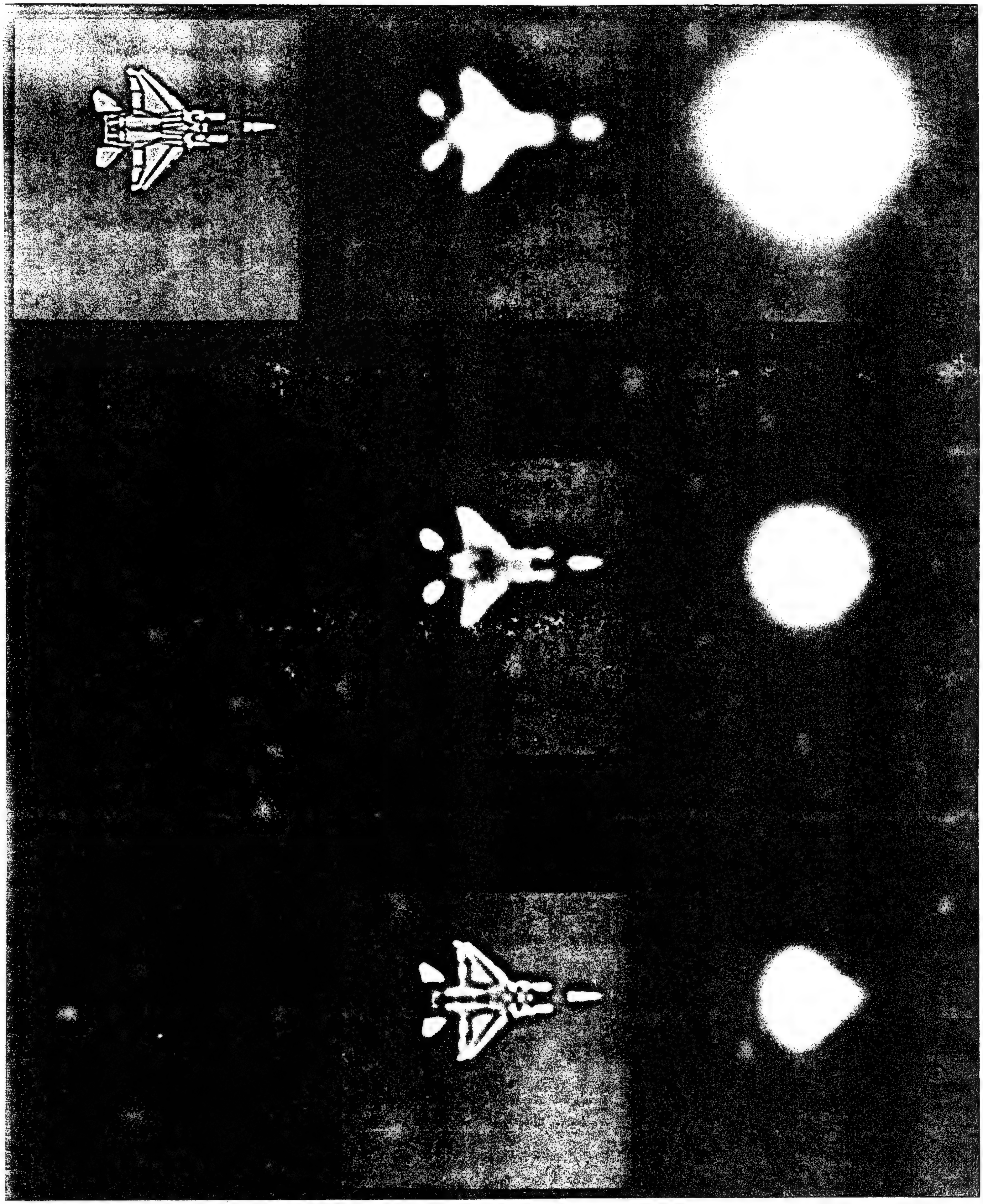

Figure 8. Bandpass Spatial Filtered Series, F-15 Fighter

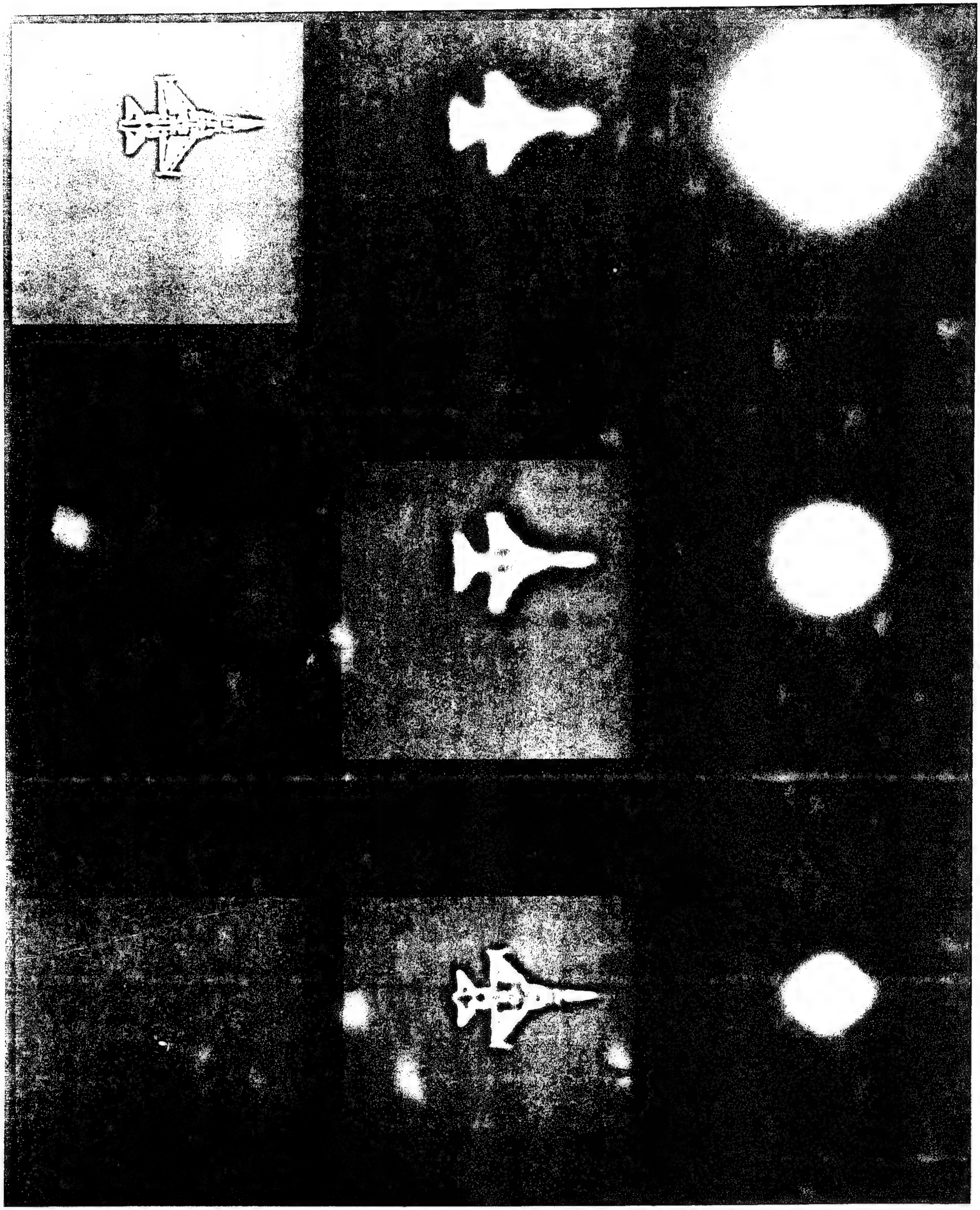

Figure 9. Bandpass Spatial Filtered Series, F-16 Fighter

SUBJECTS

Seven members assigned to the Sensor Evaluation Branch, Mission Avionics Division of the Air Force Wright Aeronautical Laboratories, Wright-Patterson Air Force Base, Ohio, served as subjects (Ss) in this study. All were male and all report 20/20 visual acuity (corrected or uncorrected). S1 through S3 were USAF trained imagery interpreters and were all familiar with the ASCC rating scale. S1 and S3 also had experience in employing other subjective image quality assessment techniques. (Kuperman, 1980, provides descriptions of some of these other techniques.) Among the trained interpreters, S3 had the least experience (12 years) and S1 had the most experience (28 years). S4 and S5 were neither trained in imagery interpretation nor experienced in image quality assessment. S6 and S7 were highly familiar with image quality assessment techniques having had on-the-job experience, as either a primary or secondary position duty, for at least 10 years. In addition, S6 had specific experience with a similar rating scale developed for SAR sensors (presented in Ward et al., 1984). All Ss participated in this study on a voluntary basis, subject to the requirement that this participation did not interfere with the performance of their mission related duties.

Because of the diversity of the Ss' backgrounds and in view of the apparent simplicity of administering the rating scale, it was decided to provide each S with an abstract task, similar to the rating scale, to provide an oppportunity for practice. The practice case, described below, was intended to serve as a familiarization with the general procedure, not as demonstration of criterion performance. (S2, the S most experienced with the rating scale, omitted the practice case.)

PRACTICE

A single set of seven practice images was prepared using the identical image processing steps employed for creating the test stimuli. The object in the practice case was a white square, 128 pixels on a side, centered in a 256 x 256 pixel black field. A practice image interpretability rating scale (contained in the appendix) was prepared using the ASCC rating scale as a

model. Level 0 on the practice scale was labeled "useless for interpretation" while level 8, the highest level employed, was labeled "perform accurate mensuration of object's dimensions." The results of the image quality practice rating sessions are summarized in Table 5.

TABLE 5. SUMMARY OF PRACTICE INTERPRETABILITY RATINGS

Filter CF (cycles/object)	Summary Statistics	
0.25	mean	0.8
	S.D.	0.7
0.50	mean	1.7
	S.D.	1.2
1.00	mean	3.8
	S.D.	1.2
2.00	mean	5.5
	S.D.	0.8
4.00	mean	6.5
	S.D.	0.5
8.00	mean	7.3
	S.D.	0.5
16.00	mean	8.0
	S.D.	0.0

CF = center frequency

S.D. = standard deviation

The monotonic increase in mean rating and the general decrease in variability with increasing center frequency strongly suggest that the Ss were properly applying the practice rating scale. Figure 10 presents the practice imagery. The unfiltered square appears in the upper left corner while the filtered versions, in order of increasing center frequency, descend from left to right. The Fourier transform of the square is presented in the lower right corner.

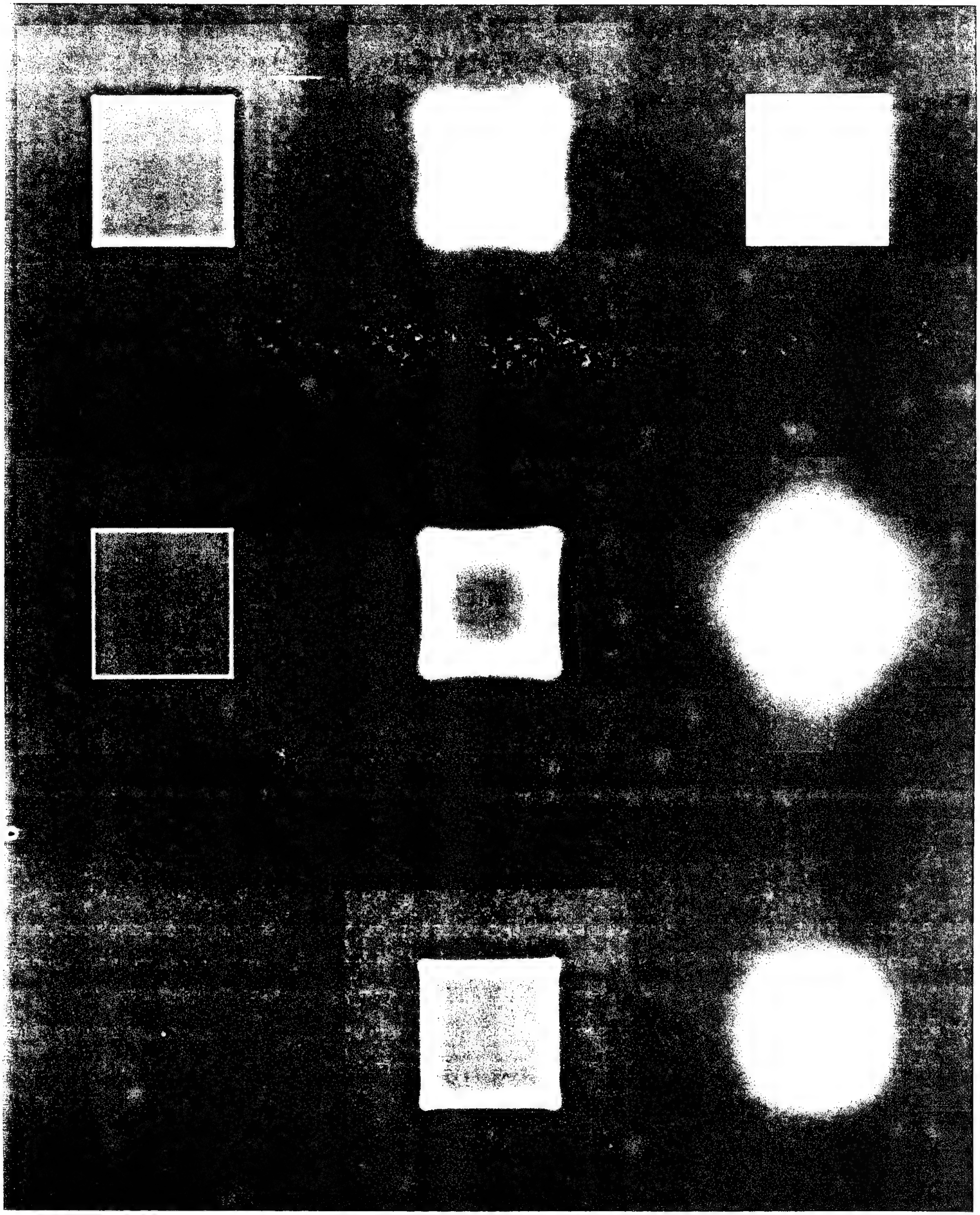

Figure 10. Spatial Filtering Series (Practice Square)

In addition to the practice rating scale, a four-point confidence scale was also employed. The Ss were required to express their confidence in their image interpretability rating using one of the descriptive phrases:

- Don't Know
- Possible
- Probable
- Certain

Table 6 presents a statistical summary of the practice confidence ratings.

TABLE 6. SUMMARY OF PRACTICE CONFIDENCE RATINGS

Filter CF (cycles/object)	Summary Statistics	
0.25	mean	3.2
	S.D.	1.0
0.50	mean	2.7
	S.D.	1.2
1.00	mean	2.8
	S.D.	0.7
2.00	mean	3.2
	S.D.	0.4
4.00	mean	3.5
	S.D.	0.5
8.00	mean	4.0
	S.D.	0.0
16.00	mean	4.0
	S.D.	0.0

CF = center frequency

S.D. = standard deviation

With the exception of the first two filters, confidence in the practice trial increased monotonically with increasing spatial frequency and, in general, the associated variability decreased. The two highest center

frequency filters produced unanimous agreement at the highest confidence report.

TASK

The experiment was self-administered by the Ss. The complete set of instructions, data recording sheets, and checklist appears in the appendix. First, the S completed the Training and Experience Questionnaire. Second, the S practiced with the filtered squares. Third, the S rated the 42 filtered images of the experiment. (Each S was presented these images in a unique random order.) Each S also provided a confidence rating (using the same four rating responses that were employed in practice) for each interpretability rating.

Section 4
ANALYSIS AND DISCUSSION

EXPERIMENTAL DESIGN

A fully repeated measures design was employed. Each of seven Ss was required to provide image interpretability and confidence ratings for each of seven filtered versions of six different U.S. tactical fighter aircraft. This resulted in a total of 294 (7 x 7 x 6) observations.

INTERPRETABILITY RATINGS

Since no evidence existed that the ASCC image interpretability rating scale possessed equal interval properties between adjacent levels of information extraction, it was necessary to convert the ratings into ranked data. This was done using the rank procedure of the SAS (1982).

It was expected that the six aircraft employed in the study would exhibit no statistically significant difference from each other and could be pooled as replicates. This hypothesis was tested using the single factor form of the SAS (1982) Analysis of Variance (ANOVA) procedure. No significant difference was, in fact, found and all subsequent analyses performed this pooling.

The remaining factors, filter and subjects, together with their interaction, pooled over the six aircraft, were tested using a two factor, ANOVA procedure (SAS, 1982). Table 7 presents a summary of this analysis. As can be seen from the table, the main effect of the Ss' performance for the seven filters (corresponding to psychovisual channels) was found to be highly significantly different. Also, the performance of the seven Ss was significantly different in an overall sense. The interaction between the filters and subjects was not found to be statistically significant.

The fact that the Ss were found to produce different image interpretability ratings was not surprising considering the differences in their training and experience levels. A post hoc test, Tukey's Honestly Significant Difference (HSD) (SAS, 1982) statistic was applied to the ranked interpretability data

TABLE 7. ANOVA SUMMARY TABLE: RANKED IMAGE INTERPRETABILITY DATA

Source	D.F.	SS	
Filters (F)	6	1728318.2	$p < 0.01$
Subjects (S)	6	106308.0	$p < 0.05$
F x S	36	121737.6	N.S.

D.F. = degrees of freedom

SS = sum of squares

N.S. = not statistically significant

in order to investigate how Ss differed from each other with respect to the interpretability scale. Figure 11 presents the results of this test. The Ss were found to be divided into four statistically distinguishable groups ($p < .05$). S5, one of the two Ss who had neither experience nor training, performed as a unique entity. [Examination of his data, pooled over filters and aircraft, revealed that, in general, he produced higher image quality ratings (and, therefore, rankings) than did any of the other Ss.] The second group was composed of S1 and S3, both trained and experienced imagery interpreters. This group was made up of S1 and S6. S6 was an experienced imagery analyst. The group that tended to produce the lowest interpretability ratings was formed by S2, S4, and S7, an interpreter, a naive S, and an analyst, respectively. (Figure 11 depicts the groupings among the Ss by the lines joining the arrows that point to the S number.) No consistent pattern is evident from this test regarding the utility of training and experience as a predictor of interpretability ratings.

It was decided to perform a post hoc computation, the omega^2 statistic (Simons, 1971), to determine the degree to which each factor in the design contributed to the total variance encountered in the resulting data. The main effect of filters accounted for 84 percent of the total variance and the main effect of subjects accounted for 5 percent of the variance. (The F x S interaction, while not significant, accounted for 6 percent of the observed variability. The remaining variance was contributed by the error term of the linear model.)

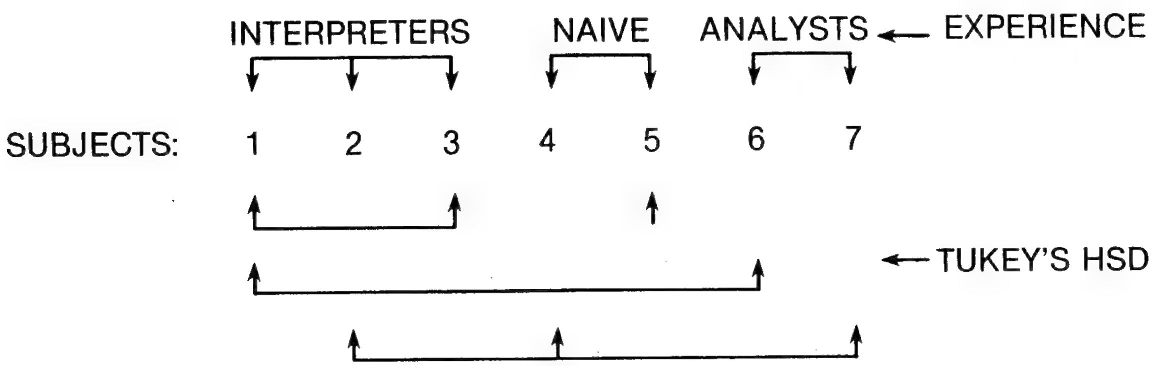

Figure 11. Post Hoc Test on Subjects (Pooled Image Interpretability Scale Data)

The nature of the highly significant difference found (in the two-factor ANOVA) for the imagery processed through the seven channel model spatial filters was also investigated using Tukey's HSD statistic (SAS, 1982) on the ranked interpretability data (pooled over aircraft). The filtered images were divided into six distinct rating groups. Images produced through the two filters having the lowest center frequencies (0.25 and 0.50 cycles per object dimension) formed one group while each of the remaining five filters (1.00, 2.00, 4.00, 8.00, and 16.00 cycles per object dimension) gave rise to imagery that was rated into a statistically separable ($p < 0.05$) image interpretability group. Figure 12 graphically depicts the means and standard deviations for the ASCC image interpretability rating scale data, pooled acrosss Ss and aircraft.

CONFIDENCE RATINGS

The interpreter confidence ratings were transformed into confidence rankings in order to achieve equal interval properties between the response levels. The SAS (1982) rank procedure was used to effect this manipulation.

A one-way ANOVA (SAS, 1982) was again applied to the factor of aircraft type (using the ranked confidence data, pooled across filters and Ss). No significant difference between aircraft types was found.

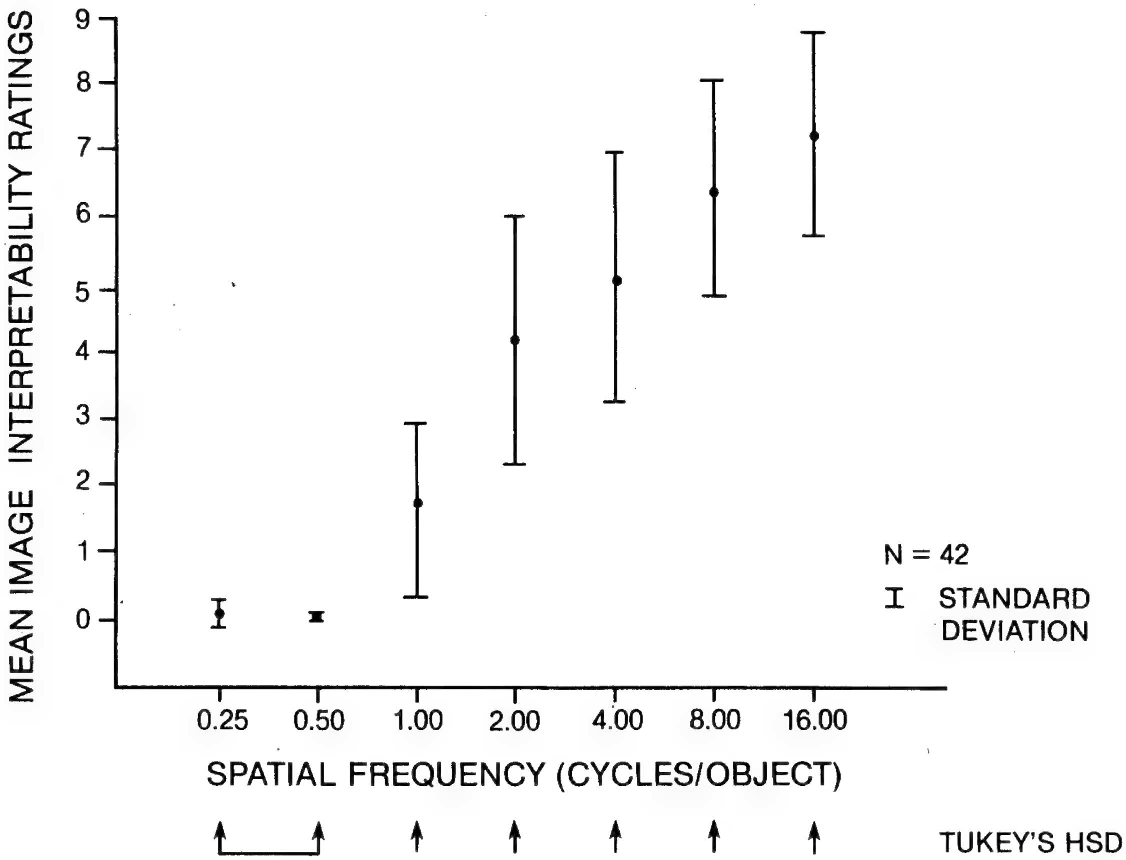

Figure 12. Means and Standard Deviations for Image Interpretability Rating Scale Data (Pooled Over Aircraft and Subjects)

The remaining factors, filters and Ss, together with their interaction, were subjected to a two-factor ANOVA (SAS, 1982). Table 8 presents the summary of this analysis. As is indicated in the table, the Ss' confidence in their interpretability ratings differed significantly on the basis of which filter had been employed in processing the imagery. Neither the main effect of Ss nor the F x S interaction was found to be statistically significant.

The omega^2 statistic (Simons, 1971) was also applied to these data to determine the contribution of each source of variation to the total observed variability in the data. The main effect of filters contributed 17 percent of the total, that of Ss accounted for 6 percent, but their interaction gave rise to 41 percent.

TABLE 8. ANOVA SUMMARY TABLE: RANKED IMAGE INTERPRETABILITY DATA

Source	D.F.	SS	
Filters (F)	6	316183.9	$p < 0.05$
Subjects (S)	6	109788.4	N.S.
F x S	36	771827.7	N.S.

D.F. = degrees of freedom

SS = sum of squares

N.S. = not significant

Tukey's HSD statistic (SAS, 1982) was applied to the ranked confidence data (pooled across Ss and aircraft). Three significantly distinct ($p < .05$) groupings of the filtered imagery were found. The first group, which gave rise to the highest levels in reported confidence, was made up of the imagery filtered at the two lowest (0.25 and 0.5 cycles per object dimension) and the two highest (8.00 and 16.00 cycles per object dimension) center frequencies. The second group, which partially overlapped the first and which gave rise to an intermediate level in mean S confidence, was made up of imagery produced by filters having the lowest (0.25) and the second and third highest (4.00 and 8.00) center frequencies (cycles per object dimension). The last group was made up of imagery created using filters with center frequencies of 1.00 and 2.00 cycles per object width; this group produced the lowest average confidence ratings. [Figure 13 presents the mean confidence rankings (normalized) in graphic form.]

REGRESSION ANALYSIS

The ANOVA on interpretability rating tested the statistical significance of the obtained rating on the spatial filter (and other factors) used to generate the test imagery. The ANOVA procedure is based on a linear model. The exact relationship between the spatial filters and the ratings can be estimated by regression. What is desired is a predictor equation for image interpretability rating based on the filter, specifically on the center frequency of the filter.

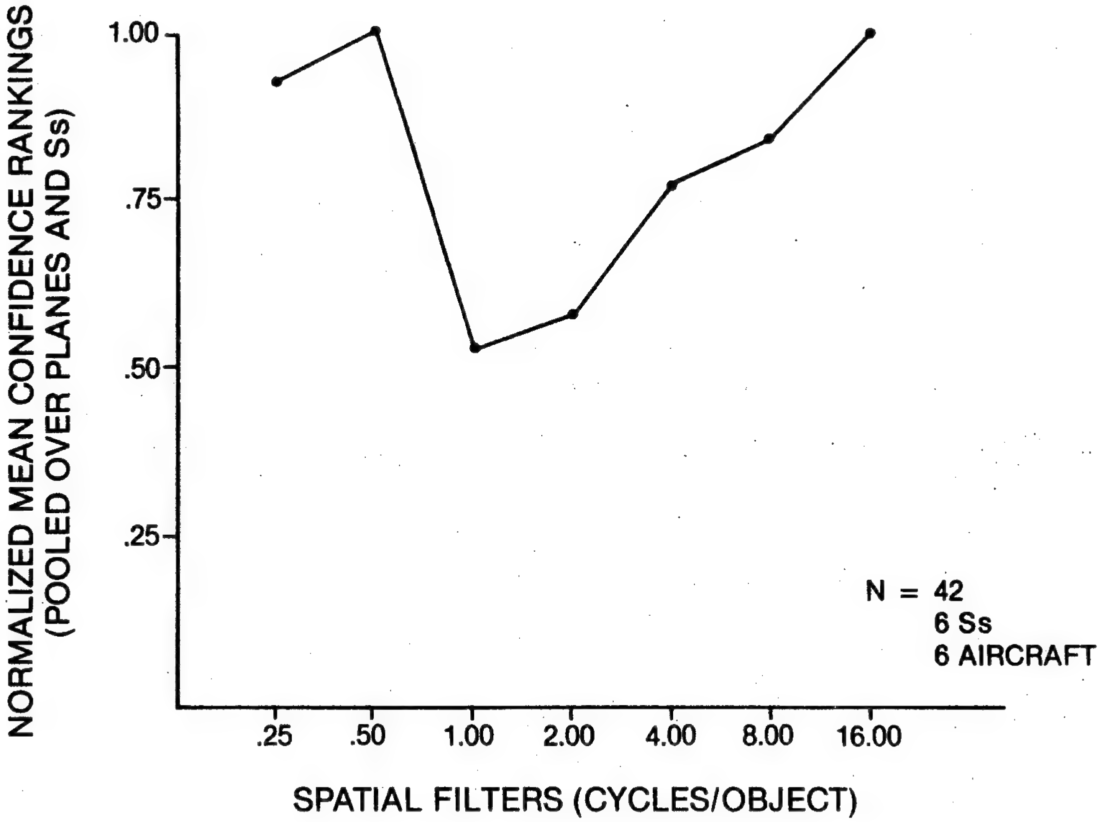

Figure 13. Mean (Normalized) Confidence Rankings

The SAS (1982) regression procedure was applied to the ratings for each observation by center frequency of the filter used to create the stimulus. The regression equation that resulted was:

$$\text{Ratings} = 0.431\ (\text{Center Frequency}) + 1.579$$

The r^2 for this equation was found to be 0.562. A scatterplot of the mean ratings (pooled over Ss) is presented in Figure 14 along with the regression line. It is apparent from the figure that the form of the regression equation is determined by the fact that the data from the three lowest center frequency and the highest center frequency filters fall, generally, below the line while the remaining means appear above it.

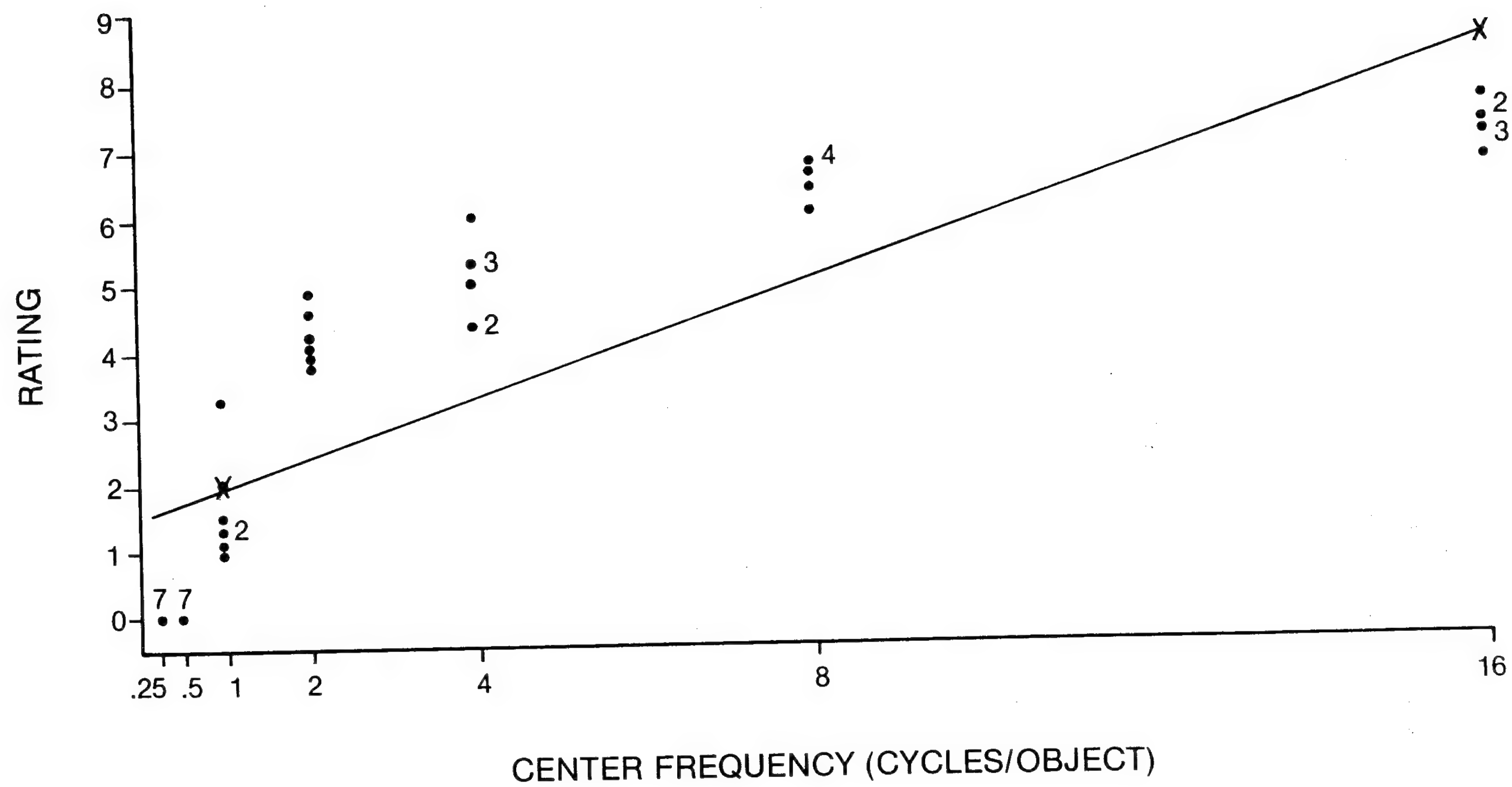

Figure 14. Scatterplot and Regression Line (Pooled Over Ss)

The r^2 of 0.562 suggests that there is linear relationship between the center frequency of the Gaussian bandpass spatial filter and the image interpretability rating to be associated with the quality of the resultant reconstruction. Table 9 presents one possible way in which the relationship between information content and digital filter center frequency may be represented. In constructing the table, the regression equation was inverted to predict center frequency as a function of rating. The levels of information that can be exploited from imagery of the rated levels shown was extracted from the ASCC rating scale descriptors.

TABLE 9. INFORMATION CONTENT

Rating	Center Frequency	Level of Information
0	--	Useless
1	--	Detect Large Aircraft
2	1.0	Count Large Aircraft (by wing type)
3	3.3	Detect/Count All Aircraft
4	5.7	Recognize Aircraft
5	7.9	Detect Presence of Alphanumerics on Wings of Large Aircraft
6	10.3	Identify Aircraft (by canopy type)
7	12.6	(No Additional Information)
8	14.9	Identify All Aircraft
9	--	(No Additional Information)

The regression analysis demonstrated a linear relationship between image interpretability rating and channel center frequency. Since the rating value is reported to correspond to the level of information available for exploitation and since these levels of information form a hierarchy, then the channel model filters may also be considered to produce a hierarchial organization of visual perception.

Section 5
CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

1. For the fighter aircraft types and digital image processing operations considered in this study, no significant difference was found between aircraft in an image interpretability rating task.

2. The channel model, when applied as digital image processing filters, produced imagery which resulted in image quality ratings (ASCC imagery interpretability scale) that were highly significantly different from each other ($p < 0.01$).

3. The rating scale employed revealed significant differences between the Ss who applied it ($p < 0.05$). No consistent pattern between the training and experience of the Ss and their ratings was observed.

4. No difference was found between aircraft types on the basis of the Ss' confidence in their interpretability ratings.

5. The channel model filters produced significant ($p < 0.05$) differences in terms of interpreter confidence reports.

6. No differences were found between Ss nor was there a significant interaction between filters and Ss for the dependent variable of confidence in interpretability rating.

7. The highest average confidence ratings were produced by the imagery created using the two highest and two lowest center frequency spatial filters (0.25, 0.50, 8.00, and 16.00 cycles per object dimension). The lowest confidence ratings came from imagery based on spatial filters of intermediate center frequency (1.00 and 2.00 cycles per object dimension).

8. A linear predictor equation was developed to relate the center frequency of the spatial filter to the interpretability rating that would be produced ($r^2 = 0.5623$). This equation, together with the descriptors associated with the ASCC rating scale, allow prediction of the information to be extracted from imagery processed with Gaussian weighted, bandpass spatial filters of the center frequencies used as the independent variables.

9. In summary, the channel model of visual perception was found to be linearly related to an operationally accepted and employed reconnaissance imagery interpretability scale. The scale was based on a hierarchy of information extraction levels. Hence, the several spatial filters which together compose the channel model, may be considered to also correspond to distinct levels of information extraction.

RECOMMENDATIONS

The channel model has been demonstrated to correlate with image quality ratings derived in the context of information extraction tasks (imagery interpretation). This finding has direct application to a variety of operational Air Force problems in which information content, stimulus complexity, image quality, display design, and/or visual information transfer are significant parameters. It is recommended that the channel model be investigated for possible application to:

1. The design of bandwidth compression algorithms which optimize the reconstructed imagery with respect to the psychovisual capabilities of the observer.

2. The assessment of display image quality.

3. The design of imagery interpretation equipment.

4. The quantification of visual fatigue experienced by imagery interpreters.

APPENDIX

TWO-DIMENSIONAL FILTERING EXPERIMENT

INSTRUCTIONS TO SUBJECTS

Purpose

Psychologists working in the area of visual perception have developed theories of vision that suggest that an image is broken down into its component parts before the brain processes the information contained in the image. This experiment is an attempt to determine what components of an image (photograph) of an aircraft contain information required by an imagery interpreter/analyst in order to carry out an image quality rating task. The results of this study will help the Air Force in designing better image interpretation equipment and methods.

All information regarding your participation in this experiment will be kept confidential. Only group data (all subjects) will be reported and your individual answers will be kept in confidence.

Please perform the experiment by yourself. Do not consult with any other subjects or compare answers.

Your part in the experiment should take between 1 and 2 hours. Please work at your own pace.

Materials

In order to carry out your part, you will need four things:

1. Subject's Packet: This large envelope contains a short questionnaire concerning your training and experience as an imagery interpreter/analyst, an answer sheet for the practice images, and an answer sheet for the test images. (A checklist is also provided to help you perform your part of the experiment in the correct sequence.)

2. Practice Imagery Packet: This large envelope contains seven small envelopes, each with a 4 x 5 inch print inside it.

3. "School Solution" Packet: This envelope contains six prints. Each print is a picture of a fighter aircraft with the designation and nickname for that aircraft given. (You will use these pictures as if they were part of an imagery interpretation key when you carry out the experiment.) A copy of the NATO Air Standardization Coordinating Committee Image Interpretability Scale (ASCC-AIR STD 101/11, 10 July 1978) is also contained in this packet.

4. Test Imagery Packet: This large envelope contains 42 small envelopes. Each small envelope (a) is marked with a code number and (b) contains a single print. These are the images that you will be rating during the experiment.

<u>Procedure</u>

Your part in the experiment is made up of the following steps:

1. Subject Packet. Take an unsealed subject packet. Remove the "Subject's Training and Experience Questionnaire." Complete the Questionnaire.

2. Practice Photos. Take the seven practice photos and the Practice Answer Sheet. Study the Interpretability Scale and Confidence Ratings on the first page of the Practice Answer Sheet. Use these ratings to rate the seven practice photos. You may rate the practice photos one at a time or all seven together. When you are finished, replace the practice photos in the Practice Photo Packet and place your Practice Answer Sheet in your Subject's Packet.

3. Experiment. Take the six "School Solution" photos and study them. Note similarities and differences between the six aircraft. (Look at the shape of the wings, the tail, the shape and position of the canopy along the fuselage, the number of engines, the shape of the nose, etc.)

Now take the 42 experiment photos and the Experiment Answer Sheet. Look at the first Image Number on the Answer Sheet. Find the photo envelope with the _same_ number marked on it. Remove the photo. (Check that the number marked on the photo is the same number as is on the envelope.) Study the photo. Using the NATO Interpretability Rating Scale, mark the interpretability level number on the Answer Sheet. Then mark your confidence rating. Replace the photo in its envelope. Go on to the next Image Number (as marked on the Answer Sheet). Continue until you have rated all 42 photos.

SUBJECT'S TRAINING AND EXPERIENCE QUESTIONNAIRE

Name ______________________________ Date ____________

Grade/Rank ______________ Organization ______________

1. Read the Instructions to Subjects. Please indicate that you have read and understand the Instructions by signing your name on the next line. (Signature) ______________________________

2. Do you wear eyeglasses or contact lenses when you work as an imagery interpreter/analyst? Yes ____ No ____

(If the answer is yes, please wear your glasses/lenses while carrying out the experiment.)

3. What training have you received as an imagery interpreter/analyst? (Please list any military schools, short courses, special certifications, and special on-the-job-training that you may have received.)

____

____

____

4. How many years have you worked as an interpreter/analyst? ____________

5. Have you ever been trained to use the NATO Rating Scale (or similar interpretability rating scales)? Yes ____ No ____

(If you have worked with similar scales, please identify them. ____________

____

6. Please list any specialized image interpretation or image analysis equipment that you have used in your work. ______________________________

____

____

____

CHECKLIST

1. Subject Questionnaire

2. Practice

   - Practice Photos
   - Practice Answer Sheet

   Rate the practice photos.

   Replace the practice photos in their packet.

   Put your Practice Answer Sheet in your Subject Packet.

3. "School Solution"

   - "School Solution" Photos

   Study them; note similarities and differences.

4. Experiment

   Arrange photo evelopes in order.

   Rate the photos.

   Refer to "School Solution" Photos (if you want to).

5. Replace Experiment Photos

   Replace "School Solution" Photos.

6. Place answer sheets and questionnaire in Subject's Packet and seal the flap.

PRACTICE ANSWER SHEET

INTERPRETABILITY

The practice photos each contain "filtered" versions of a square in the center of the photo. You are asked to estimate the level of interpretability that each practice photo would support. You are not asked to do any interpretation or analysis. Use the following ratings to rate the seven practice photos:

Level 0 -- Useless for interpretation.

Level 1 -- Detect the presence of an object.

Level 2 -- Locate the position of the object within the photograph.

Level 3 -- Determine the orientation of the object.

Level 4 -- Recognize the shape of the object with low confidence (possible).

Level 5 -- Estimate the aspect ratio (length-to-width) of the object.

Level 6 -- Recognize the shape of the object with medium confidence (probable).

Level 7 -- Recognize the shape of the object with high confidence (certainty).

Level 8 -- Perform accurate mensuration of object's dimensions.

CONFIDENCE

Please use the following descriptive words in reporting your confidence in the interpretability rating that you assign:

"Don't Know"
"Possible"
"Probable"
"Certain"

PRACTICE ANSWER SHEET

IMAGE NO.	INTERPRETABILITY	CONFIDENCE
P1	______________	______________
P2	______________	______________
P3	______________	______________
P4	______________	______________
P5	______________	______________
P6	______________	______________
P7	______________	______________

EXPERIMENT ANSWER SHEET

Use the NATO Interpretability Scale.

Use these confidence-describing terms: "Don't Know"
"Possible"
"Probable"
"Certain"

IMAGE NO.	INTERPRETABILITY	CONFIDENCE

REFERENCES

Air Standardization Coordinating Committee (ASCC), 10 July 1978, Imagery Interpretability Rating Scale, AIR STD 101/11.

AFM-200-50, 1967, Image Interpretation Handbook, Volume I.

Anonymous, November 1969, Air Reconnaissance Target Reporting Guide.

Campbell, F. W. and Robson, J. G., 1968, "Application of Fourier Analysis to the Visibility of Gratings," Journal of Physiology, 197.

Campbell, Fergus W., Nachmias, Jacob, and Jukes, John, 1970, "Spatial-Frequency Discrimination in Human Vision," Journal of the Optical Society of America, 60.

Carlson, Curtis R. and Cohen, Roger W., July 1978, Visibility of Displayed Information: Image Descriptors for Display, ONR-CR213-120-4F.

Cornsweet, Tom N., 1970, Visual Perception.

Erickson, Ronald A., October 1978, "Line Criteria in Target Acquisition with Television," Human Factors.

Ginsburg, Arthur P., December 1978, Visual Information Processing Based on Spatial Filters Constrained by Biological Data, AMRL-TR-78-129, Volumes I and II, AD A090 117.

Ginsburg, Arthur P., 1980, "Specifying Relevant Spatial Information for Image Evaluation and Display Design: An Explanation of How We See Certain Objects," Proceedings of the Society for Information Display.

Gliatti, E. L., 1977, "Modulation Transfer Analysis of Aerial Imagery," Photogrammetria, 33.

Johnson, J., October 1958, "Analysis of Image Forming Systems," Proceedings of the Image Intensifier Symposium.

Kuperman, Gilbert G., April 1980, Investigation of Photographic Image Quality Estimators, AFAMRL-TR-80-27, AD A087 805.

Kuperman, Gilbert G., Kama, William N., Fraggiotti, Joseph, and Kettlewell, John, June 1977, Research and Simulation in Support of Near Real Time/Real Time Reconnaissance RPV Systems, AMRL-TR-77-33, AD A044 598.

Kuperman, Gilbert G., Wallquist, Donald L., and Katz, Luan, October 1984, "A Digital Image Processing Facility for Human Factors Research," Proceedings of the Human Factors Society 28th Annual Meeting.

Lewis, Donald E., January 1984, Two-Dimensional Fast Fourier Transforms in Image Processing, AFAMRL-TR-84-006.

Nygaard, J. E., Slocum, G. K., Thomas, J. O., Skeen, J. R., and Woodhull, J. G., May 1964, The Measurement of Stimulus Complexity in High-Resolution Sensor Imagery, AMRL-TR-64-229.

Simons, C. W., December 1971, Considerations for the Proper Design and Interpretation of Human Factors Engineering Experiments, HAC-TR-P73-325.

Statistical Analysis Institute (SAS), 1982, SAS User's Guide: Statistics.

Snyder, Harry L., October 1974, Visual Search and Image Quality, AMRL-TR-73-114, AD A008 007.

Snyder, Harry L., Shedivy, David I., and Maddox, Michael E., February 1982, Quality Metrics of Digitally Derived Imagery and Their Relation of Interpreter Performance: III. Subjective Scaling of Hard-Copy Digital Imagery, HFL-81-3.

Task, Harry L., January 1979, An Evaluation and Comparison of Several Measures of Image Quality for Television Displays, AMRL-TR-79-7, AD A069 690.

Ward, Frank, Wilson, Denise, and Wallquist, Donald L., May 1984, Development of Color Criteria for Advanced Displays, AFAMRL-TR-84-23.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1985 - 559-006/20211